

Docket No.: 62758-063

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of : Customer Number: 20277
Haruhiko USA, et al. : Confirmation Number:
Serial No.: : Group Art Unit:
Filed: October 31, 2003 : Examiner:
For: A METHOD FOR ALLOCATING COMPUTER RESOURCE

CLAIM OF PRIORITY AND
TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT

Mail Stop Patent Application
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

In accordance with the provisions of 35 U.S.C. 119, Applicants hereby claim the priority of:

Japanese Patent Application No. 2002-369610, filed December 20, 2002

A Certified copy is submitted herewith.

Respectfully submitted,

MCDERMOTT, WILL & EMERY



Keith E. George
Registration No. 34,111

600 13th Street, N.W.
Washington, DC 20005-3096
(202) 756-8000 KEG:prg
Facsimile: (202) 756-8087
Date: October 31, 2003

日本国特許庁
JAPAN PATENT OFFICE

62758-063
USA et al.
October 31, 2003
McDermott, Will & Emery

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 Date of Application: 2002年12月20日

出願番号 Application Number: 特願2002-369610

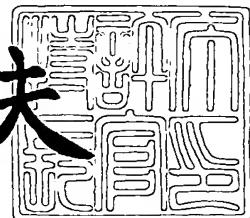
[ST. 10/C]: [JP2002-369610]

出願人 Applicant(s): 株式会社日立製作所

2003年 9月26日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願

【整理番号】 NT02P0784

【提出日】 平成14年12月20日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/177

G06F 11/34

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア事業部内

【氏名】 宇佐 治彦

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所ソフトウェア事業部内

【氏名】 内田 智斎

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 計算機資源割当方法、それを実行するための資源管理サーバおよび計算機システム

【特許請求の範囲】

【請求項 1】 計算機の資源を複数の計算機に割当てて、各々の計算機で独立してプログラムを実行する計算機システムの計算機資源割当方法において、

(1) 前記計算機の資源使用状態を収集するステップ、

(2) 前記収集したデータに基づき、各々の計算機の資源使用についての相関関係を算出するステップ、

(3) 前記収集したデータと前記算出した相関係数に基づき、各々の計算機の資源割当て値を算出し、その資源割当て値にしたがって、各々の計算機の資源割当てをおこなうステップ

を有することを特徴とする計算機資源割当方法。

【請求項 2】 前記(3)のステップが、前記収集したデータに基づいて、各々の計算機の資源使用状態を予測し、その予測した資源使用状態と前記算出した相関係数に基づき、各々の計算機の資源割当て値を算出する処理を含むことを特徴とする請求項 1 記載の計算機資源割当方法。

【請求項 3】 前記(3)のステップで資源を割当てる必要のあると判断された計算機に対して、他の計算機の資源の割当てを減じて、減じた資源をその計算機に再割り当てるに際し、

その計算機と前記相関係数の大きい計算機ほど資源の割当てを減じないことをしたことを特徴とする請求項 1 記載の計算機資源割当方法。

【請求項 4】 時間帯や各々の計算機で動作しているプログラムの特性に応じて相関係数を切り替える処理を含むことを特徴とする請求項 1 記載の計算機資源割当方法。

【請求項 5】 計算機の資源を複数の計算機に割当てて、各々の計算機で独立してプログラムを実行する計算機システムの計算機資源割当を管理するための資源管理サーバにおいて、

計算機の資源使用状態を収集する資源使用状態データ収集部と、

前記収集したデータに基づき、各々の計算機の資源使用についての相関関係を算出する相関係数算出部と、

前記収集したデータと前記算出した相関係数とに基づき、各々の計算機の資源割当て値を算出し、その資源割当て値にしたがって、各々の計算機の資源割当てをおこなう資源割当部とを有することを特徴とする資源管理サーバ。

【請求項 6】 さらに、前記収集したデータに基づいて、各々の計算機の資源使用状態を予測する資源使用予測部を有し、その予測した資源使用状態に基づき前記資源割当部が資源の割当てをおこなうことを特徴とする請求項 5 記載の資源管理サーバ。

【請求項 7】 資源割当部が、資源を割当てる必要のあると判断された計算機に対して、他の計算機の資源の割当てを減じて、減じた資源をその計算機に再割り当てるに際し、

その計算機と前記相関係数の大きい計算機ほど資源の割当てを減じないことにしたことを特徴とする請求項 5 記載の資源管理サーバ。

【請求項 8】 前記相関係数算出部が、時間帯や各々の計算機で動作しているプログラムの特性に応じて切り替えて、相関係数を算出することを特徴とする請求項 5 記載の資源管理サーバ。

【請求項 9】 計算機の資源を複数の計算機に割当てて、各々の計算機で独立してプログラムを実行する計算機システムにおいて、

この計算機システムの有する資源管理サーバは、前記計算機の資源使用状態を収集し、その収集したデータに基づき、各々の計算機の資源使用についての相関関係を算出して、前記収集したデータと前記算出した相関係数とに基づき、各々の計算機の資源割当て値を算出し、その資源割当て値を、計算機の資源割当てを制御する機構に送信し、

計算機の資源割当てを制御する機構は、その資源割当て値に基づいて、各々の計算機の資源割当てをおこなうことを特徴とする計算機システム。

【請求項 10】 前記資源管理サーバは、資源を割当てる必要のあると判断した計算機に対して、他の計算機の資源を、その資源を割当てる必要のあると判断した計算機の資源として移動させるときに、

その計算機と前記相関係数の大きい計算機ほど資源の移動をさせないようにし

その計算機と前記相関係数小さい計算機ほど資源の移動をさせるようにすることを特徴とする請求項 9 記載の計算機システム。

【請求項 11】 前記資源を割当てる計算機が、複数の計算機上に構成されたことを特徴とする請求項 9 記載の計算機システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、計算機資源割当方法に係り、複数の仮想計算機への資源を動的に割り当てるにあたって、資源の割当てを最適化して、各々の仮想計算機の相関から計算機資源の配分を理想的におこなう計算機資源割当方法に関する。

【0002】

【従来の技術】

仮想計算機システムにおいては、CPU（命令プロセッサ）、メモリ（主記憶）、および、チャネルなどの物理計算機が有している資源を、ハイパバイザが論理的に分割し、複数の仮想計算機 LPAR に割当てる。仮想計算機 LPAR (Logical Partition) は、実在する物理計算機の資源を論理的に分割している仮想的な計算機である。

【0003】

仮想計算機システムについては、例えば、特許文献 1 の従来技術の項で紹介されている。

【0004】

また、仮想計算機システムに割当てられたメモリの構成を動的に変更する方法としては、特許文献 2 に開示されている。

【0005】

【特許文献 1】

米国特許第 4 564 903 号明細書

【特許文献 2】

特開平6-348584号公報

【0006】

【発明が解決しようとする課題】

仮想計算機システムは、一台のハードウェアとしての計算機で複数のOSを同時に実行することができ、用途によっては非常に有用なシステムであるということができる。

【0007】

そして、仮想計算機システムにおいては、資源の割当では、負荷の高い仮想計算機LPARほど、多く割り当てることが望ましく、資源を動的に変更する機能が必要となってくる。

【0008】

従来の仮想計算機システムの資源の割当てにおいては、自システム（自LPAR）の負荷の変化に基づいておこなっているため、他システム（他LPAR）と連携した複合システムの場合に、他システムの負荷の変化を想定して資源の割当てを変更することはできなかった。そのため、資源の割当てを変更しても、近い将来に他システムで性能不足が発生するおそれがあり、他システムと調整して資源の割当てを変更しても、他システムに性能不足が波及しないように資源を配分することは困難であった。

【0009】

例えば、仮想計算機システムの各LPARにおいて、インターネットのWEBサーバ、データベースサーバ、開発用のテストサーバをそれぞれ運用しており、WEBサーバの負荷が増大すると近い将来にデータベースサーバの負荷が増大するといった相関関係が見られる場合であっても、WEBサーバの負荷が増大した時点では、近い将来データベースサーバの性能が不足し得ることを想定した資源の再割当をおこなう仕組みは、従来では提供されておらず、データベースサーバの性能が不足した時点で再び資源の再割当をおこなう必要があるという問題点があった。

【0010】

本発明は、上記問題点を解決するためになされたもので、その目的は、複数の

仮想計算機への資源の割当てを動的に再配分するにあたって、資源の割当てを最適化して、各々の仮想計算機の相関から計算機資源の配分を理想的におこなうようにして、近い将来に他の仮想計算機の性能不足が発生しにくいように、各仮想計算機に割当てられた資源を配分することを可能にする計算機資源割当方法を提供することにある。

【0011】

【課題を解決するための手段】

上記問題点を解決するために、本発明の仮想計算機システムの計算機資源割当方法は以下のようにした。資源管理サーバが、仮想計算機L P A Rの資源の使用状態を収集し、収集したデータに基づき、資源の使用状態を予測する。また、過去の仮想計算機L P A Rの実行履歴により、各々の仮想計算機L P A Rの資源の使用状態についての相関関係を算出する。

【0012】

そして、予測値と算出した相関係数に基づき、各々の仮想計算機L P A Rの資源割当て値を算出し、その資源割当て値にしたがって、各々の仮想計算機L P A Rの資源割当てをおこなう。

【0013】

このときに、ある仮想計算機L P A Rにおいて資源の割当不足が予測されるとき、その仮想計算機L P A Rとの相関係数が小さい仮想計算機L P A Rに割当てていた資源を優先的に資源の割当が不足した仮想計算機L P A Rへ割当て、その仮想計算機L P A Rとの相関係数が大きい仮想計算機L P A R（近い将来に性能不足が発生しやすい傾向があるL P A R）に割当てていた資源はなるべく減じないことにする。

【0014】

これは、二つの仮想計算機L P A Rの相関係数が大きい場合、一方の仮想計算機L P A Rが使用する資源が増加すると他方の仮想計算機L P A Rが使用する資源も同時に、あるいは、近い将来に増加する傾向がある。すなわち、ある仮想計算機L P A Rにおいて資源不足が予測されるとき、その仮想計算機L P A Rとの相関係数が大きい仮想計算機L P A Rは、近い将来に性能不足が発生しやすい傾

向があるためである。

【0015】

このようにすることにより、近い将来に各仮想計算機L P A Rの性能不足が発生しにくいように、各仮想計算機L P A Rに割当てられた資源を再配分することが可能になる。

【0016】

すなわち、上記の手段によって、資源管理サーバが各L P A Rの資源を管理するシステムにおいて、あるL P A Rの資源の割当不足が予測されたときに、各L P A R間の相関関係に基づき、資源の割当不足が予測されたL P A Rとの相関関係が低いL P A Rから優先的にC P U割当率、およびメモリ割当量を減じることにより、効率よく各L P A Rに資源を再割当てすることが可能になる。

【0017】

また、各仮想計算機L P A Rを、複数の物理計算機上に構成し、資源管理サーバが資源の管理を複数の物理計算機上にわたっておこなえるようにする。

【0018】

このようにすれば、各L P A Rが複数の物理計算機にある場合に、これらの物理計算機の合計の資源割当の上限が定められた設定であっても、設定の範囲内で効率よく各物理計算機の資源の割当を増減しながら各L P A Rの資源の割当を再配分することが可能になる。

【0019】

【発明の実施の形態】

以下、本発明に係る各実施形態を、図1ないし図14を用いて説明する。

【0020】

【仮想計算機システムの構成】

先ず、図1を用いて本発明に係る計算機資源割当方法をおこなう仮想計算機システムの構成について説明する。

図1は、本発明に係る計算機資源割当方法をおこなう仮想計算機システムの構成図である。

【0021】

本発明の仮想計算機システムは、物理計算機121と資源管理サーバ101が、ネットワーク131により、接続された構成である。

【0022】

ここで、物理計算機121と言うのは、仮想計算機と対比した語であり、ハードウェアとしての計算機に、論理的な仮想計算機が構築されることを意味している。

【0023】

資源管理サーバは、物理計算機121上に構築される仮想計算機LPAR122上に割当てる資源を管理して、適切な資源配分をおこなうために指示を与えるサーバである。

【0024】

資源管理サーバ101は、機能モジュールとして、資源使用状態収集部102、相関係数算出部103、資源使用予測部104、資源不足検出部105、資源割当決定部106を持ち、データテーブルとしては、資源使用状態テーブル107、相関係数テーブル108、資源使用予測テーブル109、資源割当設定テーブル110、および、資源割当情報テーブル111を備えている。

【0025】

物理計算機121は、複数の仮想計算機LPAR122が構築され、独立して動作することができる。また、物理計算機121上のCPU、メモリが各仮想計算機LPAR122に割り当てられ、見かけ上は、各々の仮想計算機LPAR122が、CPU124、メモリ125を有しているように見ることができる。また、仮想計算機LPAR122は、資源使用測定部123を有し、その仮想計算機LPAR122の資源の使用に関するデータを測定している。

【0026】

ハイパバイザ126は、物理計算機121を論理的に分割し、複数の仮想計算機LPAR122を構成するための制御機能であり、各仮想計算機LPAR122に資源を割り当てるための資源割当部127を有している。

【0027】

仮想計算機LPAR122の資源使用測定部123は、定期的にLPAR12

2の資源の使用状態に関するデータ、すなわち、C P U 1 2 4の使用率、および、メモリ125の使用量を測定し、資源管理サーバ101の資源使用状態収集部102へ測定した資源の使用状態に関するデータを送信する。資源使用状態収集部102は、受け取った資源の使用状態に関するデータを収集し、資源使用状態テーブル107、および資源割当情報テーブル111へ格納する。

【0028】

次に、相関係数算出部103は、資源使用状態テーブル107を使用して各L P A Rの相関係数を算出し、相関係数テーブル108へ格納する。相関係数とは、各仮想計算機L P A R 1 2 2が、動作時に他の仮想計算機L P A R 1 2 2とどのような資源の使用状態の相関を持って動作するかを表す指数であり、これについては後に説明する。

【0029】

次に、資源使用予測部104は、前記資源使用状態収集部102がデータを収集するたびに、前記資源使用状態テーブル107を使用して各L P A Rの、動作状態における資源の使用状態を予測して、資源使用予測テーブル109へ格納する。

【0030】

次に、資源不足検出部105において、格納した資源使用予測テーブル109に基づき、各L P A Rの資源が不足するかどうか判定する。資源が不足する場合には、資源割当決定部106において、資源の再割当の配分を決定し、資源割当情報テーブル111へ決定した資源の割当に関する情報を格納し、さらに、資源割当情報テーブル111のデータを、ハイパバイザ126の資源割当部127へ送信する。資源割当部127は、その配分情報に従って、仮想計算機L P A R 1 2 2に対するC P U 1 2 4、および、メモリ125の割当配分を変更する。

【0031】

なお、本実施形態では、計算機資源は、C P Uとメモリを例にして説明するが、その他の計算機資源でもよい。例えば、仮想計算機L P A R 1 2 2のディスクの数、チャネルの数などのI／Oに関する資源でもよい。

【0032】

〔計算機資源割当方法のためのデータ構造〕

次に、図2ないし図6を用いて本発明に係る計算機資源割当方法のためのデータ構造について説明する。

図2は、資源使用状態テーブル107のテーブル構造を示す図である。

図3は、相関係数テーブル108のテーブル構造を示す図である。

図4は、資源使用予測テーブル109のテーブル構造を示す図である。

図5は、資源割当設定テーブル110のテーブル構造を示す図である。

図6は、資源割当情報テーブル111のテーブル構造を示す図である。

【0033】

資源使用状態テーブル107は、仮想計算機LPAR122ごとに用意され、各資源の状態を時系列で格納するためのテーブルであり、図2に示されるように、LPAR番号201を有し、さらに、CPU使用率203、および、メモリ使用量204を時刻202についての時系列で格納される。

【0034】

CPU使用率203には、時刻202に示される時刻にその仮想計算機LPAR122が、物理計算機121のCPUを実際に使用した時間の割合を百分率(%)で示した値が格納される。例えば、10:25から10:30までの5分間に、LPAR1がCPUを合計2分間使用した場合、CPU使用率は2分/5分×100=40%である。メモリ使用量204には、その仮想計算機LPAR122が実際に使用したメモリの量が格納される。

【0035】

この資源使用状態テーブル107には、このように資源使用状態収集部102が各LPAR122から収集した資源の使用状態に関するデータが時系列で格納され、相関係数算出部103における相関係数の算出、資源使用予測部104における資源の使用状態の予測のために使用される。

【0036】

相関係数テーブル108は、仮想計算機LPAR122の資源の使用状態の実績から仮想計算機LPAR122間の資源の使用状態の相関を表した相関係数を格納するためのテーブルであり、図3に示されるように、LPAR番号301ご

とに、各仮想計算機L P A R 1 2 2の全てのL P A R 3 0 2、3 0 3、3 0 4との組み合わせについての相関係数が格納される。

【0037】

相関係数とは、任意の2つのL P A Rの資源の使用状態の相関関係を示す値である。L P A R iとL P A R jの相関係数を k_{ij} とすると、 $0 \leq k_{ij} \leq 1$ であり、 $k_{ij} = 0$ のときは両者の資源の使用状態には相関関係がなく、 $k_{ij} = 1$ のときは、両者の性能には密接な相関関係があるものとして定義する。相関係数 k_{ij} が大きい（1に近い）場合には、L P A R iが使用する資源が増加するとL P A R jが使用する資源も同時に、あるいは、近い将来に増加する傾向があることに注意しておく。また、相関係数 k_{ij} が小さい（0に近い）場合には、L P A R iが使用する資源は、L P A R jが使用する資源の増減に影響されず、無関係に増減する傾向があることに注意しておく。

【0038】

この相関係数テーブル108には、相関係数算出部103が資源使用状態テーブル107に基づいて算出した相関係数が格納され、資源割当決定部106において仮想計算機L P A R 1 2 2に対する資源の割当てのために使用される。

【0039】

は資源使用予測テーブル109は、各仮想計算機L P A R 1 2 2ごとの資源の使用状態を予測した値を格納するためのテーブルであり、図4に示されるように、L P A R番号401ごとに予測C P U使用率402、および、予測メモリ使用量403が格納される。

【0040】

この資源使用予測テーブル109には、資源使用予測部104が資源使用状態テーブル107に基づいて算出した予測データが格納される。例えば、資源使用状態テーブル107に5分間隔でデータが格納され、10:30のデータまで格納されたとき、資源使用予測部104は次のタイミング、すなわち10:35に予測されるデータを算出し、資源使用予測テーブル109へ前記予測されるデータを格納する。

【0041】

資源割当設定テーブル110は、仮想計算機LPAR122ごとに、資源の割当てのための範囲を規定するためのテーブルであり、図5に示されるように、LPAR番号501ごとに、その仮想計算機LPAR122が契約しているCPU割当率の最大値502、最小値503、および、メモリ割当量の最大値504、最小値505が格納される。

【0042】

資源割当設定テーブル110には、予め、各LPARの資源割当てのための最大値、最小値を設定しておき、その値が変更されるときには更新される。

【0043】

CPU割当率とは、物理計算機121が有するCPUをその仮想計算機LPAR122に割当てている時間を百分率（%）で示したものである。例えば、5分間に30秒間だけLPAR1へCPUを割当てている場合には、CPU割当率は10%である。CPU割当率とCPU使用率は異なる値であり、同じ時間帯ではCPU割当率 \geq CPU使用率である。例えば、前記30秒間の割当て（CPU割当率は10%）のうち、実際に、LPAR1がCPUを使用した時間が15秒間であれば、CPU使用率は5%である。同様に、メモリ割当率とは、物理計算機が有するメモリを該LPARに割当てた量である。同じ時間帯ではメモリ割当率 \geq メモリ使用量である。

【0044】

資源割当情報テーブル111は、各仮想計算機LPAR122に対する資源の割当てを決定するために使用されるテーブルであり、LPAR番号601ごとに、CPU割当率602、およびメモリ割当量603を有する。

【0045】

変更前の資源割当情報テーブル111（図6（a））には、資源使用状態収集部102が各LPARから収集した資源の使用状態の情報が格納され、資源不足検出部105、および、資源割当決定部106によって資源の割当てを決定するために使用される。

【0046】

資源割当決定部106によって、決定された資源の割当ての情報は、再び、資

源割当情報テーブル111に格納される。そして、この変更された資源割当情報テーブル111（図6（b））の値は、ハイパバイザ126の資源割当部127へ送信される。

【0047】

〔計算機資源割当方法のための処理〕

次に、図7ないし図13を用いて本発明に係る計算機資源割当方法のための処理について説明する。

【0048】

最初に、図7により本発明の計算機資源割当方法の処理の概要について説明する。

図7は、本発明に係る計算機資源割当方法の処理を示すゼネラルチャートである。

【0049】

先ず、資源使用状態収集処理では、資源管理サーバ101の資源使用状態収集部102は、各仮想計算機LPAR122の資源使用状態を収集して、その仮想計算機LPAR122の資源使用状態テーブル107に格納する（S701）。

【0050】

次に、相関係数算出処理では、資源管理サーバ101の相関係数算出部102は、資源使用状態テーブル107を参照して、仮想計算機LPAR122間の相関係数を求めて、相関係数テーブル108に格納する（S702）。

【0051】

次に、資源使用予測処理では、資源管理サーバ101の資源使用予測部104は、資源使用状態テーブル107を参照して、仮想計算機LPAR122間の資源の使用状態を予測して、資源使用予測テーブル109に格納する（S703）。

【0052】

次に、資源割当決定処理では、資源管理サーバ101の資源割当決定部106は、資源の割当ての変更をおこなう仮想計算機LPAR122を決定して、新たな資源割当ての配分を求めて、それを資源割当テーブル111に格納して、ハイ

パーバイザ126に送信する（S704）。

【0053】

以下、各処理の詳細について説明する。

【0054】

先ず、図8により資源使用状態収集処理について説明する。

図8は、資源使用状態収集処理を示すフローチャートである。

【0055】

先ず、資源使用状態収集部102は、仮想計算機LPAR122から、例えば、以下の表1に示されるような資源の使用状態のデータ001を収集する（S801）。

【0056】

【表1】

表1

LPARから収集する資源使用状態データ

	時刻	CPU 使用率	メモリ 使用量	CPU 割当率	メモリ 割当量
LPAR 1	10:30	40%	640MB	40%	800MB
LPAR 2	10:30	10%	60MB	30%	300MB
LPAR 3	10:30	20%	260MB	30%	500MB

そして、資源の使用状態のデータ001に含まれる時刻003、CPU使用率004、メモリ使用量005を、それぞれ資源使用状態テーブル107の時刻202、CPU使用率203、メモリ使用量004に格納する（S802）。

【0057】

次に、資源の使用状態のデータ001に含まれるC P U割当率006、メモリ割当量007を、それぞれ資源割当テーブル111のC P U割当率602、メモリ割当量603に格納する（S803）。

【0058】

次に、図9により相関係数算出処理について説明する。

図9は、相関係数算出処理を示すフローチャートである。

【0059】

先ず、相関係数算出部103は、相関係数算出部103は、資源使用状態テーブル107から各仮想計算機L P A R 1 2 2の資源使用状態のデータを取得する（S901）。次に、各仮想計算機L P A R 1 2 2間の相関係数を算出する。

【0060】

相関係数は、図3の相関係数テーブルの見られるように、L P A R番号の全ての組み合わせについてそれぞれ算出する（S902）。例えば、L P A Rがn個あるときは、 $n \times n$ 個の組み合わせについて相関係数を算出する。L P A R_iの資源使用状態テーブル107からC P U使用率203、あるいは、メモリ使用量204を時刻202についての時系列で取り出し、これらをベクトルとして表現し、 $p_i = (p_{i1}, p_{i2}, \dots, p_{it})$ としたとき、L P A R_iとL P A R_jの相関係数 k_{ij} はベクトルの内積およびベクトル長を使用して、以下の（式1）で求めることができる。

【0061】

【数1】

数1

$$k_{ij} = \frac{(p_i \cdot p_j)}{|p_i| \cdot |p_j|}$$

$$= \frac{\sum_t p_{it} \cdot p_{jt}}{\sqrt{\sum_t p_{it}^2} \sqrt{\sum_t p_{jt}^2}} \cdots (\text{式1})$$

このようにして、相関係数は、CPU使用率、およびメモリ使用量のそれぞれについて算出することができる。そして、算出した相関係数を相関係数テーブル108へ格納する（S903）。相関係数は、CPU使用率、およびメモリ使用量のそれぞれについて格納することができる。またいずれか一方のみ、あるいは両者の平均値を格納することもできる。

【0062】

各LPAR上で動作しているプログラムは時間帯によってオンライン運用やバッチ運用といったように大幅に特性が異なるため、前記相関係数の算出に使用する資源使用状態のデータを時間帯によって切り分けることにより、時間帯ごとに最適な相関係数を算出することができる。また、新たな運用を開始する場合には、相関係数の算出に使用するための資源の使用状態に関するデータが未整備であることが考えられるため、新たな運用をおこなうための仮想計算機LPAR122についての相関係数を、別の手段により計算するなり、予測するなどして、入力してテーブルに格納することもできる。

【0063】

次に、図10により資源使用予測処理について説明する。

図10は、資源使用予測処理を示すフローチャートである。

【0064】

先ず、資源使用状態テーブル107から各仮想計算機LPAR122のCPU使用率203、および、メモリ使用量204を時刻202についての時系列で取得する（S1001）。そして、LPARそれぞれについて、前記取得した資源使用状態データに基づいて資源の使用状態を予測する（S1002）。資源使用状態の予測においては、例えば、最近の過去m回の資源使用状態データを滑らかな曲線、あるいは、直線で結ぶm-1次関数を利用するという技法により、次に資源使用状態データが送られてくるタイミングに相当する時刻の資源使用状態を導き出すことができる。

【0065】

資源使用状態の予測はCPU使用率、およびメモリ使用量のそれぞれについて算出する。

【0066】

次に、前記予測した値を資源使用予測テーブル109の予測CPU使用率402、および、予測メモリ使用量403に格納する（S1003）。

【0067】

次に、図11により資源割当決定処理について説明する。

図11は、資源割当決定処理を示すフローチャートである。

【0068】

先ず、資源割当情報テーブル111から各仮想計算機LPAR122のCPU割当率602、および、メモリ割当量603を取得する（S1101）。

【0069】

次に、資源使用予測テーブル109から各仮想計算機LPAR122の予測CPU使用率402、および、予測メモリ使用量403を取得する（S1102）。

。

【0070】

次に、資源割当設定テーブル110から各仮想計算機LPAR122の最大CPU割当率502、および、最大メモリ割当量504を取得する（S1103）

。

【0071】

そして、S1104からS1107の処理について、LPAR番号 $i = 1, 2, 3$ のそれぞれについて繰り返す。

【0072】

LPAR i のCPU、メモリのそれぞれについて、資源の不足が予測され、かつ、CPUやメモリの割当てを増強できる場合、すなわち、条件式「割当値 < 予測値、かつ、割当値 < 最大割当値」が満たされる場合には、S1106に進み、資源割当配分決定処理をおこない、前記条件式が満たされない場合には、S1107に進み、資源割当配分決定処理をおこなう（S1105）。資源割当配分決定処理は、サブルーチンであり、次に詳細に説明する。

【0073】

LPAR番号 $i = 1, 2, 3$ のそれぞれについての処理が終了している場合は、S1108に進む（S1107）。

【0074】

最後に、資源割当情報テーブル111に格納されているデータをネットワーク131を経由してハイパバイザ126の資源割当部127へ送信する（S1108）。

【0075】

次に、図12により資源割当配分決定処理について説明する。

図12は、資源割当配分決定処理を示すフローチャートである。

【0076】

この処理は、図11のS1105でコールされる処理であり、資源の不足が予測された LPAR i が発生した場合に、他の仮想計算機 LPAR122 から資源を移動させて、LPAR i と他仮想計算機 LPAR122との相関関係に応じて資源の不足が予測された LPAR i に再割当てをおこなう処理である。

【0077】

先ず、資源使用予測テーブル109から各仮想計算機 LPAR122 の予測値を取得する（S1201）。予測値とは、予測CPU使用率402、および予測

メモリ使用量403のうち、S1105の判定において資源割当の不足が検出されたものである。

【0078】

次に、資源割当情報テーブル111から各仮想計算機LPAR122の割当値を取得する（S1202）。割当値とは、CPU割当率602、および、メモリ割当量603のうち、S1105の判定において、資源割当の不足が検出されたものである。

【0079】

次に、 $L P A R_j$ の予測される資源割当不足値「 $d_i = \text{予測値} - \text{割当値}$ 」を算出する（S1203）。

【0080】

次に、 $L P A R_j$ ($j = 1, 2, 3$) の予測される未使用予測値「 $s_j = \text{割当値} - \text{予測値}$ 」を算出する（S1204）。 $s_j < 0$ のときは $s_j = 0$ とする。

【0081】

次に、相関係数テーブル108から $L P A R_i$ と各 $L P A R_j$ との相関係数 k_{ij} を取得する。

【0082】

S1206からステップ1208は、 $L P A R$ 番号 $j = 1, 2, 3$ について繰り返す処理である。

【0083】

前記算出した d_i 、 s_j 、および、前記取得した k_{ij} をもとに $L P A R_j$ ($j = 1, 2, 3$) の割当値を変更する（S1207）。 $L P A R_j$ の割当値の変更分 Δ_j は以下の（式2）によって算出することができる。

【0084】

【数2】

数2

$$\Delta_j = \frac{d_i \cdot s_i \cdot (1 - k_{ij})}{\sum_l s_l \cdot (1 - k_{il})} \quad \cdots \text{(式2)}$$

ここで、前記変更分 Δ_j が s_j よりも大きいときは $\Delta_j = s_j$ とする。また、 Δ_j は、前記の（式2）に限らず、相関係数 k_{ij} に基づく任意の配分方法で決定することができる。前記算出した Δ_j を前記割当値から減じた値を、資源割当情報テーブル111のCPU割当率602、ないし、メモリ割当量603へ格納する（S1209）。

【0085】

ここで、図3、図4、および、図6に示す数値を用い、LPAR1のCPU資源の割当が不足し、他のLPARに割当てていたCPU資源を減じて、LPAR1へ割当てるケースにおける具体例を説明する。

【0086】

LPAR1のCPU資源不足値は、予測値 $402 = 50\%$ 、割当値 $602 = 40\%$ であるため、LPAR1の資源不足値は「 $d_1 = 10\%$ 」である。また、各LPAR*i*の予測されるCPUの未使用予測値 s_i は「 $s_1 = 0\%、s_2 = 30\% - 10\% = 20\%、s_3 = 30\% - 20\% = 10\%$ 」である。したがって、各LPAR*i*から減じるCPU割当率 Δ_i は $\Delta_1 = 0\%、\Delta_2 = 8.57\% \approx 9\%、\Delta_3 = 1.43\% \approx 1\%$ である。すなわち、LPAR2から $\Delta_2 = 9\%$ 、LPAR3から $\Delta_3 = 1\%$ のCPU資源を削減してLPAR1へ「 $\Delta_2 + \Delta_3 = 10\%$ 」のCPU資源を割当てることができる。

【0087】

そして、 $\Delta_2、\Delta_3$ により資源の再配分をおこなうことにより、LPAR1、LPAR2、LPAR3の新たな構成は、LPAR1のCPU割当率 $= 40\% + \Delta_2 + \Delta_3 = 50\%$ 、LPAR2のCPU割当率 $= 30\% - \Delta_2 = 21\%$ 、LPAR3

3のCPU割当率 = 30% - Δ₃ = 29%となる。

【0088】

このケースでは、LPAR1とLPAR3の相関係数が大きい（1に近い）ため、LPAR1のCPU資源が不足すると近い将来にLPAR3のCPU資源も不足しやすい傾向があるが、上記算出したように、LPAR1の資源の割当てが不足するときに、LPAR1との相関関係の低いLPAR2のCPU割当率をΔ₂ (= 9%) によって多く減じ、LPAR1との相関関係の高いLPAR3は、近い将来の資源の割当て不足に備えて、Δ₃ (= 1%) で示される値しか割当率を減少させないので、LPAR3の資源の割当てをあまり減らさないで済むことになる。

【0089】

なお、上記の実施形態での説明では、資源の割当を調節するために、資源使用状態のデータから予測値を求め、それにより、仮想計算機LPAR122の資源の再配分を調節する方法について述べてきた。しかしながら、資源使用状態のデータから予測値を求めずとも、直接に、図2の資源使用状態テーブルのデータを参照して、資源の割当をおこなうための仮想計算機LPAR122とその資源の配分の割合を求めるこにしてもよい。

【0090】

次に、図13により資源使用測定処理について説明する。

図13は、資源使用測定処理を示すフローチャートである。

【0091】

資源使用測定処理は、各仮想計算機LPAR122上の資源使用測定部123により、一定時間間隔でシステムが停止するまでおこなわれる。

【0092】

S1301からS1304は、一定の時間間隔でシステムが停止するまで繰り返し動作する。

【0093】

先ず、各仮想計算機LPAR122のCPU124のCPU使用率とCPU割当率、および、メモリ125のメモリ使用量とメモリ割当量を測定する（S13

02)。

【0094】

次に、前記測定した資源使用データ、L P A R番号、および、時刻を前記表1の資源使用データ001の形式によって、資源管理サーバ101の資源使用状態収集部102へ送信する(S1303)。資源使用状態収集部102は資源使用データ001を受信すると、図7に示した計算機資源割当方法の処理を開始する。

【0095】

〔他の実施形態〕

以下、本発明に係る計算機資源割当方法をおこなう仮想計算機システムの他の構成について説明する。

図14は、本発明に係る計算機資源割当方法をおこなう仮想計算機システムの他の構成図である。

【0096】

本実施形態では、物理計算機1403は、第一の実施形態と同様に複数の仮想計算機L P A R 1404を有する。資源管理サーバ1401、および、L P A R 1404は、ネットワーク1402により接続され、資源管理サーバ1401が、C P Uやメモリなどの資源の管理をおこない資源の割当配分の指示を各仮想計算機L P A R 1404におこなうのも同様である。

【0097】

本実施形態では、各仮想計算機L P A R 1404が、複数の物理計算機1403にわたって、構成されていることが異なっている。そして、資源管理サーバ1401は、複数の物理計算機1403にまたがる各仮想計算機L P A R 1404の資源の割当てを調整することが可能である。

【0098】

すなわち、仮想計算機L P A R 1404が、異なる物理計算機にある場合であっても、上記図1を用いて説明した仮想計算機システム全く同じ方法により、複数ある物理計算機の合計性能を増減することなく、各L P A RのC P U割当率、およびメモリ割当量を再配分することができる。このように、複数ある物理計算

機のC P U資源、およびメモリ容量を各々変更することができ、資源の割当ての合計の上限が定められた設定で計算機システムを運用している場合にも、このシステムにより、合計の資源の割当てを一定に保ちながら、各物理計算機の資源の割当てを増減し、各仮想計算機L P A R 1 4 0 4へ有効に資源を割当てることが可能になる。

【0099】

〔本実施形態の応用〕

各仮想計算機L P A RにおいてインターネットのW E Bサーバ、データベースサーバ、開発用テストサーバといった異なる業務を運用しているシステムがあり、W E Bサーバの負荷が増大すると、近い将来にデータベースサーバの負荷も増大するが、開発用テストサーバの負荷の増減はW E Bサーバの負荷の増減とは無関係である、といった相関関係が見られる場合を想定する。

【0100】

この場合に、W E Bサーバの負荷が増大し、資源の割当不足が予測された時点において、相関関係の低い開発用テストサーバのC P U割当率およびメモリ割当率をより多く減じることにする。このようにすれば、相関関係の強いデータベースの負荷が近い将来増大した場合に、再び、各仮想計算機L P A Rの資源の割当率を変更し直さなければならないという事態が発生することを予防することが可能になる。

【0101】

さらに、本実施形態では、一台の物理計算機の資源を複数の仮想計算機L P A Rに割当てる例について説明したが、資源を割当てる計算機は、物理計算機であっても同様に本発明は適用することができる。すなわち、資源管理サーバを置き、物理計算機の要求に応じて、C P U資源やメモリなどの割当てをおこなう場合にも本発明の資源割当ての手法を用いることにより、資源の割当てを最適化して、各々の計算機の相関から計算機資源の配分を理想的におこなう計算機システムを構築することができる。

【0102】

〔発明の効果〕

本発明によれば、複数の仮想計算機への資源の割当てを動的に再配分するにあたって、資源の割当てを最適化して、各々の仮想計算機の相関から計算機資源の配分を理想的におこなうようにして、近い将来に他の仮想計算機の性能不足が発生しにくいように、各仮想計算機に割当てられた資源を配分することを可能にする計算機資源割当方法を提供することができる。

【図面の簡単な説明】

【図 1】

本発明に係る計算機資源割当方法をおこなう仮想計算機システムの構成図である。

【図 2】

資源使用状態テーブル 107 のテーブル構造を示す図である。

【図 3】

相関係数テーブル 108 のテーブル構造を示す図である。

【図 4】

資源使用予測テーブル 109 のテーブル構造を示す図である。

【図 5】

資源割当設定テーブル 110 のテーブル構造を示す図である。

【図 6】

資源割当情報テーブル 111 のテーブル構造を示す図である。

【図 7】

本発明に係る計算機資源割当方法の処理を示すゼネラルチャートである。

【図 8】

資源使用状態収集処理を示すフローチャートである。

【図 9】

相関係数算出処理を示すフローチャートである。

【図 10】

資源使用予測処理を示すフローチャートである。

【図 11】

資源割当決定処理を示すフローチャートである。

【図12】

資源割当配分決定処理を示すフローチャートである。

【図13】

資源使用測定処理を示すフローチャートである。

【図14】

本発明に係る計算機資源割当方法をおこなう仮想計算機システムの他の構成図である。

【符号の説明】

101…資源管理サーバ

102…資源使用状態収集部

103…相関係数算出部

104…資源使用予測部

105…資源不足検出部

106…資源割当決定部

107…資源使用状態テーブル

108…相関係数テーブル

109…資源使用予測テーブル

110…資源割当設定テーブル

111…資源割当情報テーブル

121…物理計算機

122…仮想計算機L P A R

123…資源使用測定部

124…C P U

125…メモリ

126…ハイパバイザ

127…構成変更部

131…ネットワーク

1401…資源管理サーバ

1402…ネットワーク

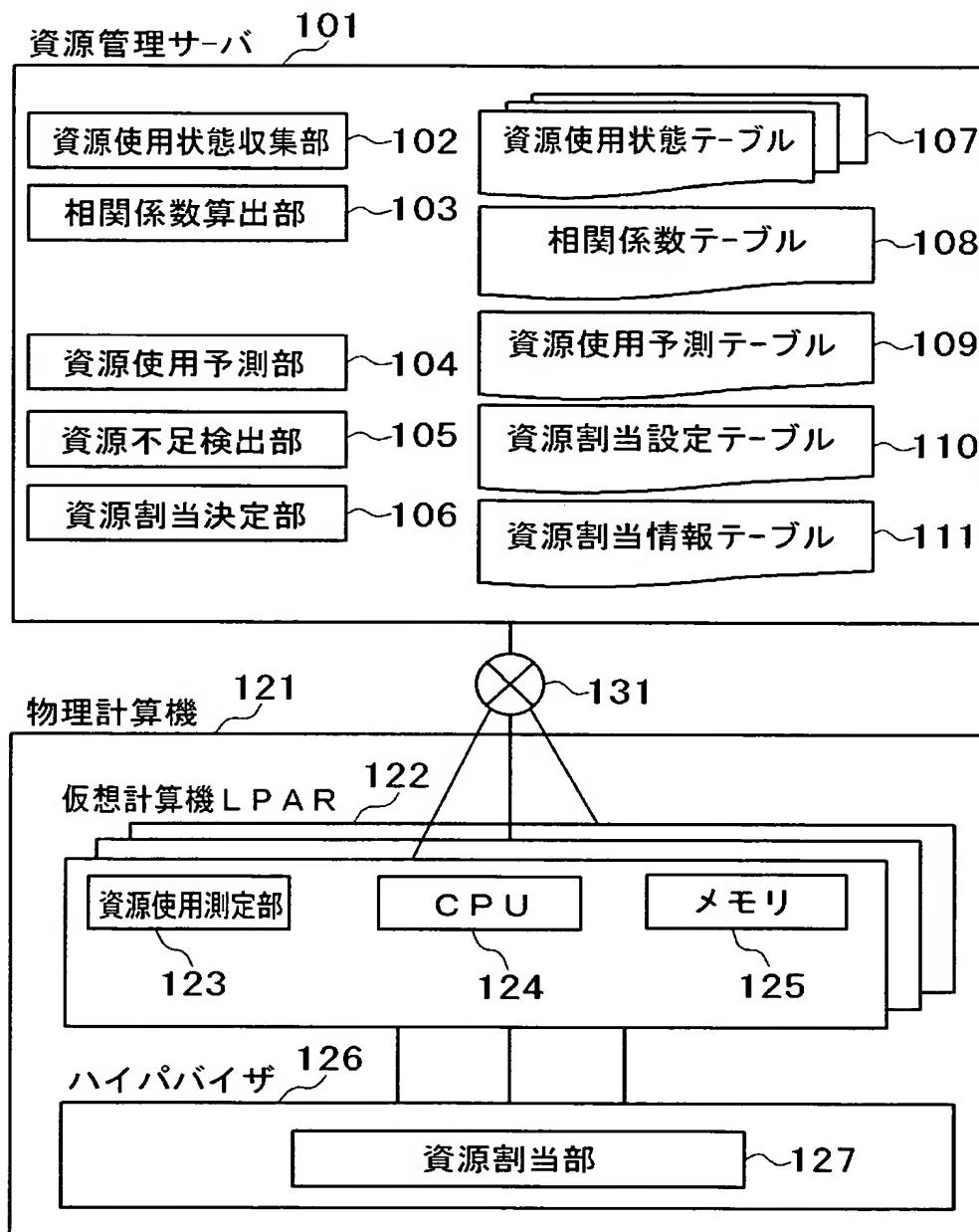
1403…物理計算機

1404…仮想計算機L P A R。

【書類名】 図面

【図 1】

図 1



【図 2】

図 2

資源使用状態テーブル

107

LPAR 3		
LPAR 2		
LPAR 1		
時刻	CPU 使用率	メモリ 使用量
...
10:20	30%	300MB
10:25	25%	280MB
10:30	40%	640MB

201 202 203 204

10:30までのデータ

【図 3】

図 3

相関係数テーブル

108

	LPAR 1	LPAR 2	LPAR 3
LPAR 1	1	0.1	0.7
LPAR 2	0.1	1	0.3
LPAR 3	0.7	0.3	1

301 302 303 304

【図4】

図4

資源使用予測テーブル

	予測CPU使用率	予測メモリ使用量	109 10:35に予測されるデータ
LPAR1	50%	1200MB	
LPAR2	10%	60MB	
LPAR3	20%	280MB	

401

402

403

【図5】

図5

資源割当設定テーブル

	CPU割当率		メモリ割当量		110
	最大	最小	最大	最小	
LPAR1	70%	10%	1600MB	256MB	
LPAR2	30%	10%	512MB	64MB	
LPAR3	70%	10%	1500MB	128MB	

501

502

503

504

505

【図6】

図6

資源割当情報テーブル(変更前) 111

	CPU割当率	メモリ割当量
LPAR1	40%	1600MB
LPAR2	30%	512MB
LPAR3	30%	1500MB

↓ 資源割当決定部により割当を変更

資源割当情報テーブル(変更後) 111

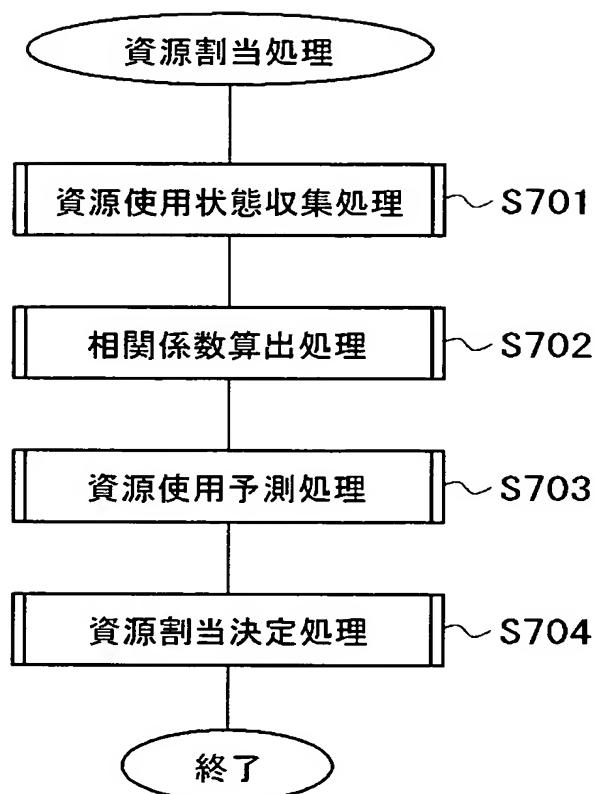
	CPU割当率	メモリ割当量
LPAR1	50%	1200MB
LPAR2	21%	70MB
LPAR3	29%	330MB

10:30の割当

新たな割当

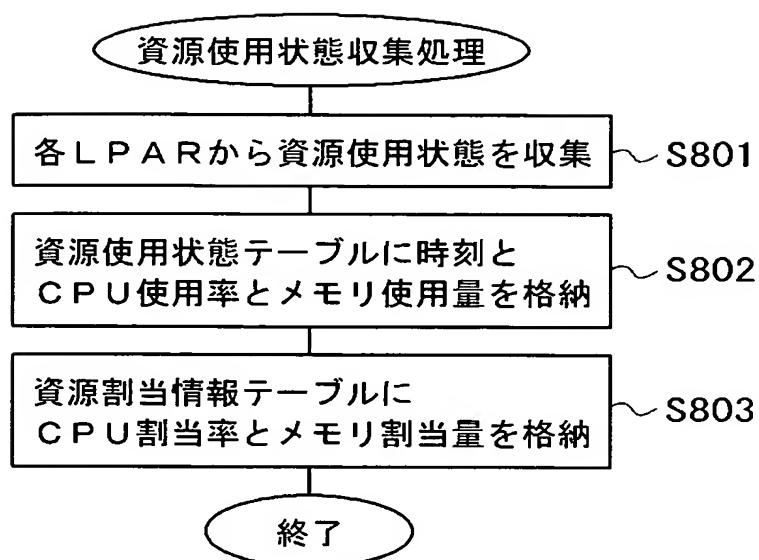
【図 7】

図 7



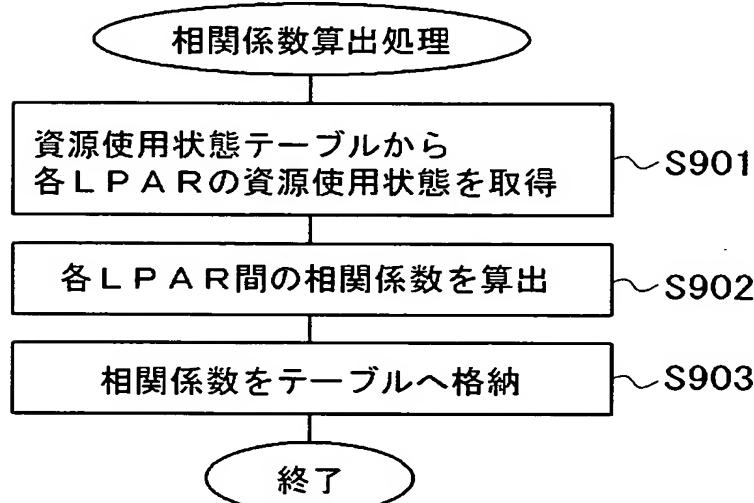
【図8】

図8



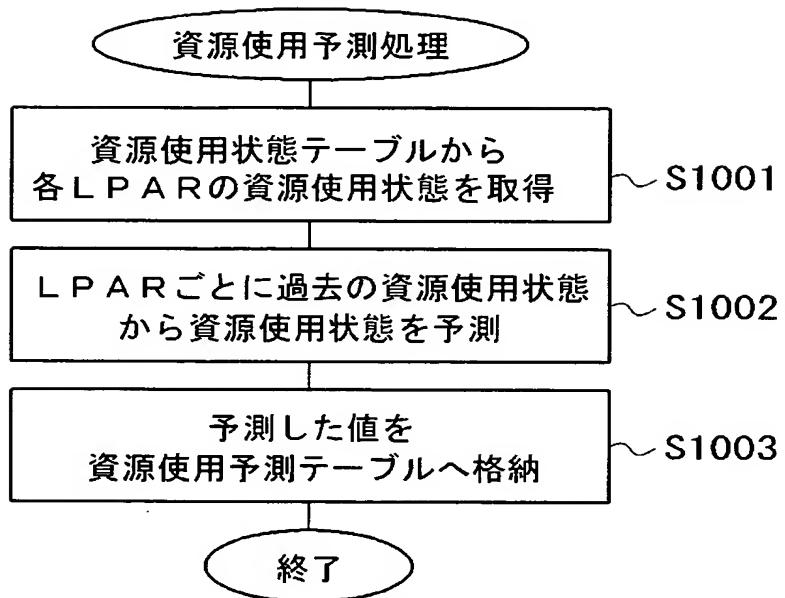
【図9】

図9



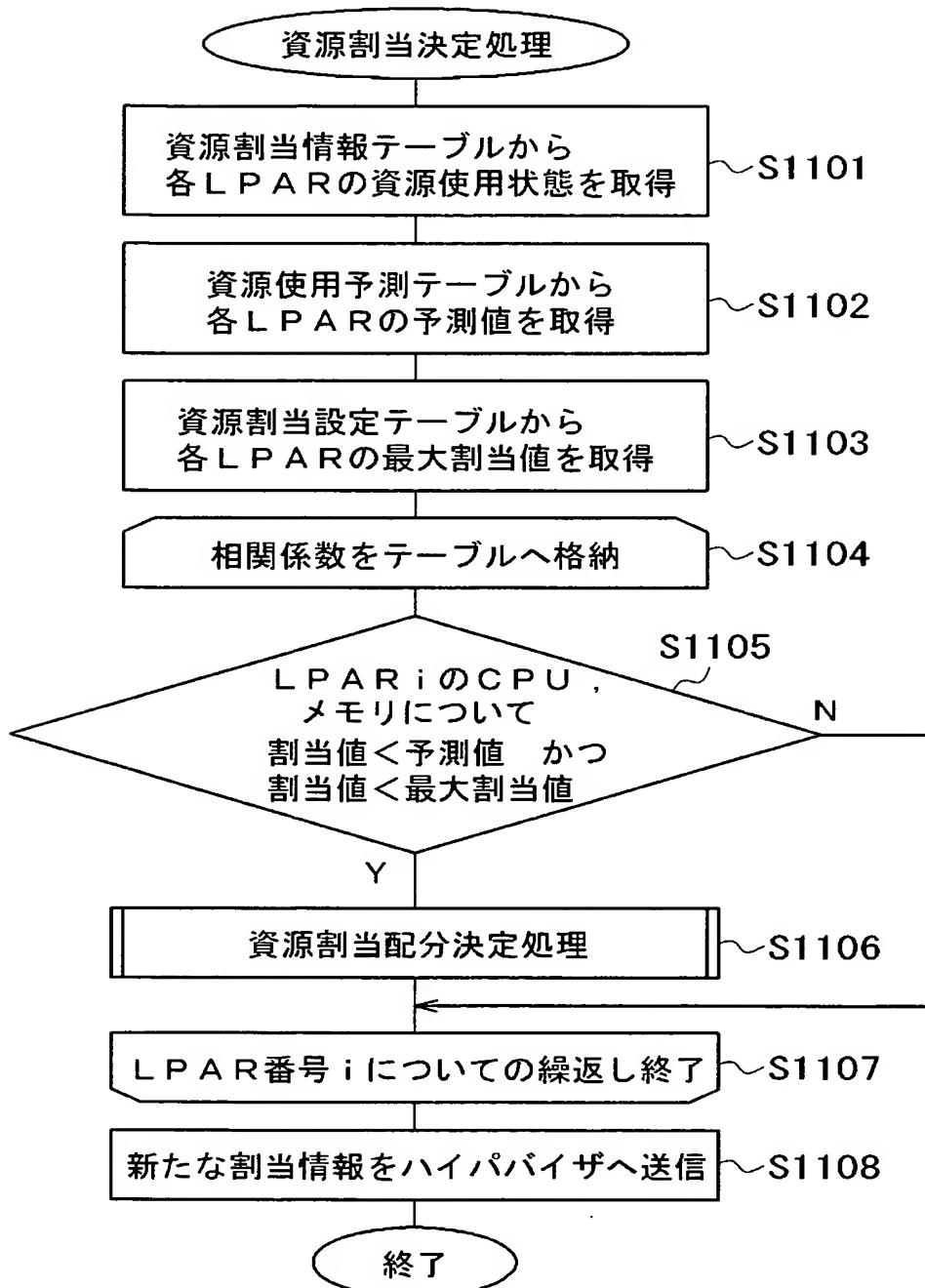
【図10】

図10



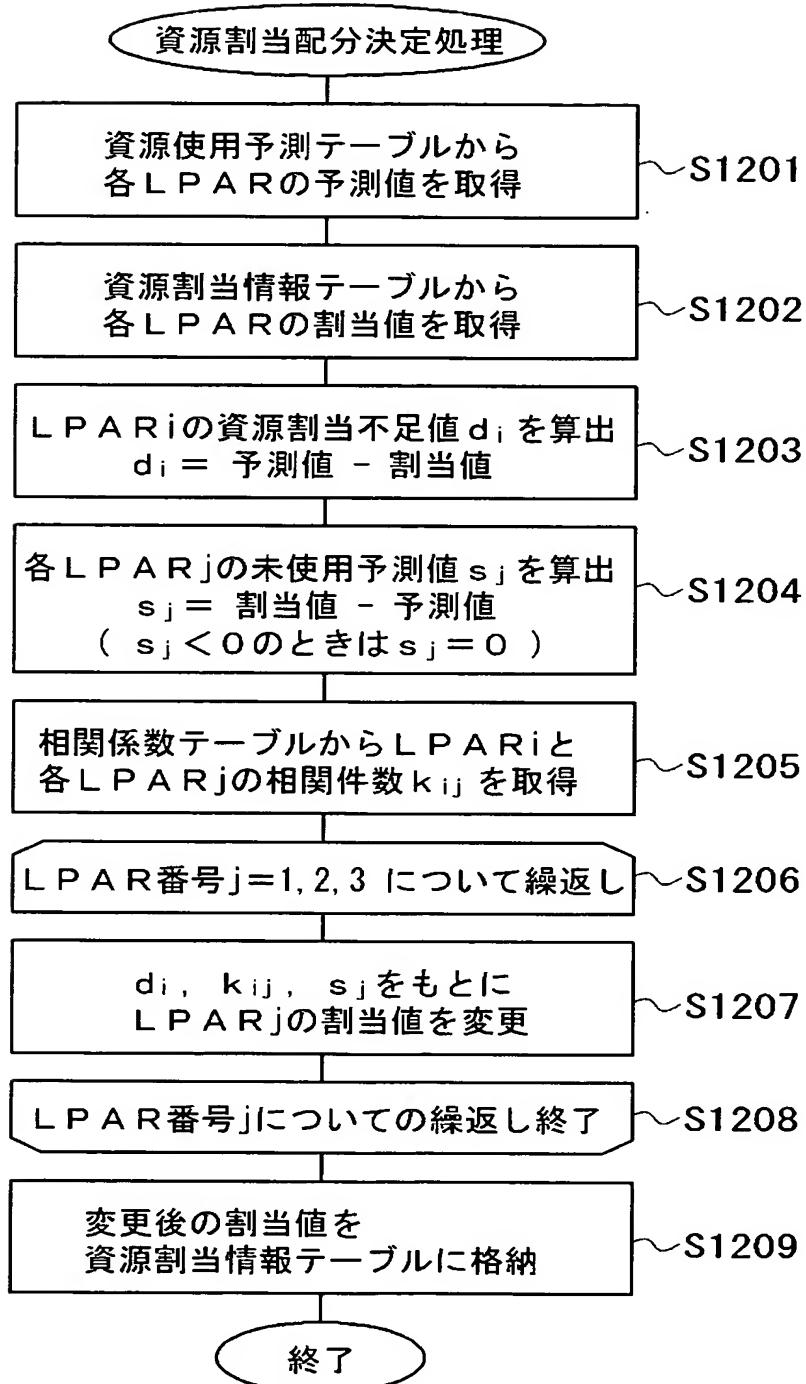
【図11】

図11



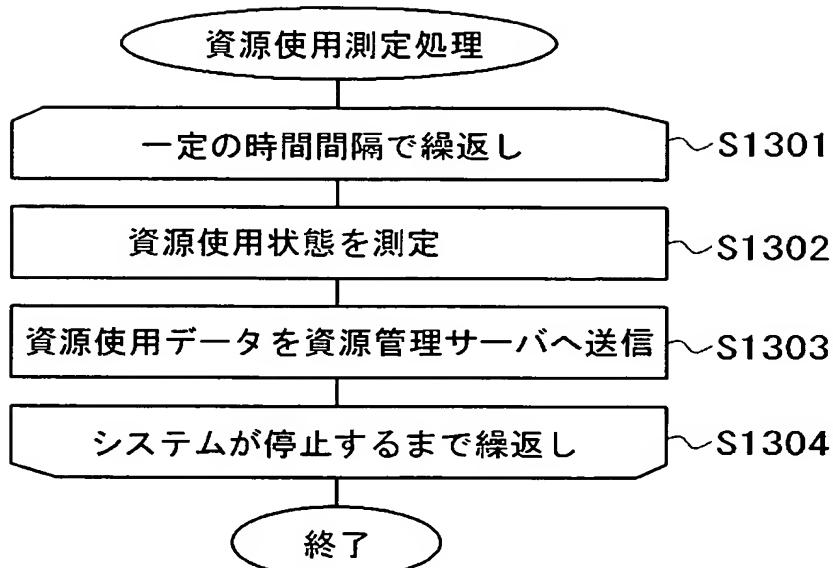
【図12】

図12



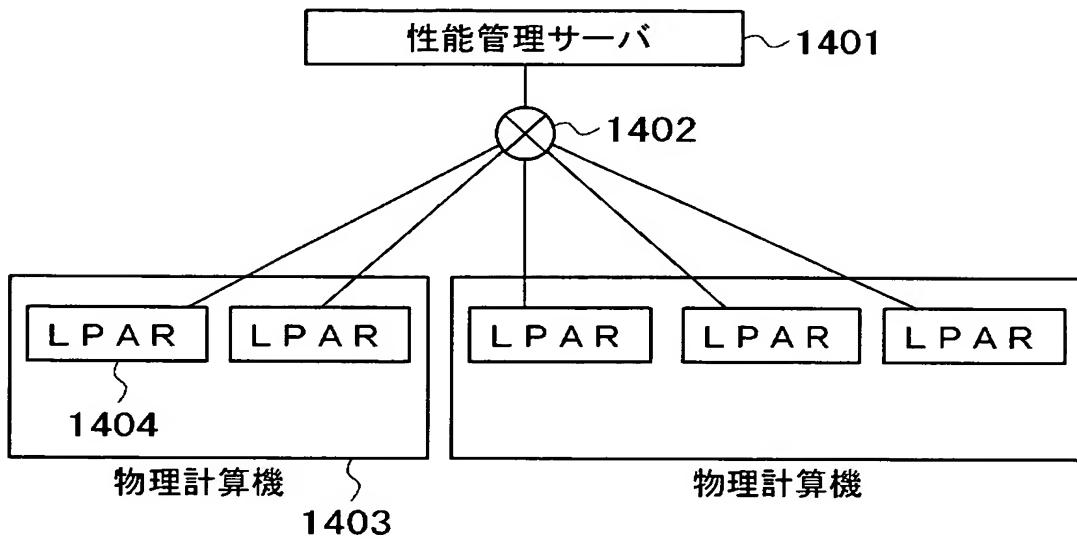
【図13】

図13



【図14】

図14



【書類名】 要約書

【要約】

【課題】複数の仮想計算機への資源の割当てを動的に再配分するにあたって、資源の割当てを最適化して、近い将来に他のL P A Rの性能不足が発生しにくいように、各L P A Rに割当てられた資源を配分する。

【解決手段】資源管理サーバが、仮想計算機L P A Rの資源の使用状態を収集し、収集したデータに基づき、資源の使用状態を予測する。また、過去の仮想計算機L P A Rの実行履歴により、各々の仮想計算機L P A Rの資源の使用状態についての相関関係を算出する。そして、予測値と算出した相関係数に基づき、各々の仮想計算機L P A Rの資源割当て値を算出し、その資源割当て値にしたがつて、各々の仮想計算機L P A Rの資源割当てをおこなう。

【選択図】 図1

特願2002-369610

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所